

# Set-Theoretic Measures as Evolutionary Fitness Criteria in Nonlinear System Identification <sup>\*</sup>

Jinyao Yan and J.R. Deller, Jr.

*Department of Electrical and Computer Engineering, and  
NSF BEACON Research Center, Michigan State University,  
E. Lansing, MI 48824, USA (e-mail: {yanjinya, deller}@egr.msu.edu).*

**Abstract:** The paper reports a method for identification of parametric models that are linear and time-invariant in parameters, but arbitrarily nonlinear in signals. Set-bounding solutions are exploited to simultaneously identify and parametrize the model structure. Measures of set-solution quality are used as fitness measures in evolving the structure. Experiments verify effective NARX and NARMAX identification in complex unmodeled disturbances.

© 2015, IFAC (International Federation of Automatic Control) Hosting by Elsevier Ltd. All rights reserved.

**Keywords:** Nonlinear systems, parameter estimation, model selection, set-membership identification, evolutionary algorithm.

## 1. INTRODUCTION

Contemporary modeling problems increasingly require solutions that transcend conventional linear, time-invariant (LTI) system approaches. Nonlinear and/or time-varying models are difficult around which to design, analyze, and compute solutions. In spite of significant innovation and effort among the system identification (SYSID) research community [e.g., (Billings and Zhu, 1994; Milanese et al., 1996; Wigren, 2006; Keesman, 2002)], nonlinear model identification poses many unanswered questions in contrast with the treasury of theoretical and practical LTI knowledge available to the SYSID practitioner. In this paper, we adopt one approach to accounting for non-LTI system properties, without entirely abandoning the advantages of LTI identification. This compromise is achieved using models that are *LTI in parameters* (LTiP), with nonlinearities accounted for in the signal processing by the system. LTiP nonlinear models are versatile and only marginally more difficult to work with than systems with purely LTI structure.

Any LTiP-based identification procedure must select effective model terms (regressor signals) and estimate the parameters from inputs and output observations in the presence of noise that is generally correlated or with nonlinear dependencies. An approach by Sjöberg et al. (1995) uses an initial model with a broad collection of nonlinear terms, then prunes (*backward elimination*) the insignificant components. This approach is known to cause numerical and computational problems (Billings, 2013). Sjöberg et al. (1995) also suggest a residual-based selection approach in which terms are selected sequentially according to a measure of goodness of fit (*forward selection*). An example of this approach is the FROLS algorithm (Chen et al., 1991; Billings and Zhu, 1994) which is based on

the orthogonal least squares (OLS) estimator. In order to achieve unbiased parameter estimation in the presence of colored or more complex noise, FROLS must repeatedly refit the noise model. Moreover, the forward selection is greedy by only adding one term at a time.

A significant body of work by SYSID researchers has used of parameter-set-bounding approaches, often called *set-membership (SM) algorithms*. This research can be partitioned into two categories. The first involves explicitly nonlinear system approaches, many resulting in algorithms of high computational complexity. Efforts of this type are reported in the seminal literature on set-based methods,<sup>1</sup> and the comprehensive study by Milanese and Novara (2004) involving noise and functional gradient bounds provides a more recent example of this type with a clear contrast to the second category. The second partition of set-based approaches uses more conventional LTI models (ARX-like with various noise models) in which LTiP results are implicit, but not the focus of the work.<sup>2</sup> The latter work has not featured nonlinearities because, with a predetermined model structure, the nonlinear aspects of a LTiP model are present only as numerical observations with a cumulative effect in the residuals. Accordingly, results are very nonspecific to the effects of particular nonlinearities, but at the same time, are quite sensitive to them. In this paper, however, the LTiP structure is of very deliberate concern. Contrary to previous work, the nonlinear structure of the model does not remain static in the identification process.

Two lines of reasoning underlie the methods in this paper. First, it is difficult to assess the effects of a predetermined,

<sup>\*</sup> This work was supported in part by the U.S. National Science Foundation under Cooperative Agreement DBI-0939454. Any opinions, conclusions or recommendations expressed are those of the authors and do not necessarily reflect the views of the NSF.

<sup>1</sup> The literature in this area is vast. The following citations, e.g., provide extensive reference lists: (Milanese et al., 1996; Milanese and Vicino, 1996; Deller et al., 1993; Walter and Piet-Lahanier, 1990).

<sup>2</sup> In addition to the citations in footnote 1, the paper by Deller et al. (1994) provides a unified view of the “second partition” methods. Fogel and Huang (1982) published the seminal work of this type employing linear input-output models.

static nonlinear structure in a LTIIP model. The most challenging problem in an identification problem is ordinarily the determination of the model structure. Second, in spite of many interesting and beneficial features of set-bounding algorithms, the information inherent in *set* solutions has not been extensively researched for potential exploitation. This work uses measures of model quality reflected in set solutions for guidance in the selection of models. This is accomplished through evolutionary strategies with fitness measures derived from the set solutions. Accordingly, technique simultaneously solves the model structure identification and parameter-estimation problems, in the presence of unknown noise scenarios.

## 2. IDENTIFICATION FRAMEWORK

In the present approach, the parameter estimation task uses set-theoretic analysis of the data to deduce feasible sets of solutions in light of certain model assumptions. SM algorithms, as described in Section 1, provide sets of feasible parameter vectors rather than the single point estimates of conventional estimators. This is achieved through successive refinements of an initial solution set, consistent with *a priori* constraints on the signal or system model. In the present work, measurable set solution properties are used to assess the viability of nonlinear regressor functions that compete for “survival” as components of the model (Yan et al., 2013, 2014). Specifically, we describe an evolutionary approach to the selection of nonlinear regressors. A very significant advantage of the use of a SM algorithm is the lack of need for assumptions about stationarity or distributional characteristics of the disturbances.

Consider a single-input–single-output discrete-time system with input  $x$  and output  $y$ , each typically assumed to belong to some well-behaved space,  $\mathfrak{X} \subset \mathbb{R}^Z$  (e.g.,  $\ell_2$ ). Let  $\mathbb{F}_\theta : \mathfrak{X} \rightarrow \mathfrak{X}$  denote a continuous system operator mapping  $x$  to  $y$ , which is parametrized by a real vector  $\theta \in \mathfrak{P} \subset \mathbb{R}^Q$ ,

$$y = \mathbb{F}_\theta(x). \quad (1)$$

The system is said to be *linear-in-parameters* if, for any  $x \in \mathfrak{X}$ , for all  $\theta, \theta' \in \mathfrak{P}$ , and for all  $\alpha, \alpha' \in \mathbb{R}$ ,

$$\mathbb{F}_{\alpha\theta + \alpha'\theta'}(x) = \alpha\mathbb{F}_\theta(x) + \alpha'\mathbb{F}_{\theta'}(x). \quad (2)$$

The internal processing of the system is based on a subset of a *candidate set* of nonlinear regressor functions,  $\Xi_\varphi = \{\varphi_q\}$ , of size  $|\Xi_\varphi|$ . Each regressor is a mapping  $\varphi_q : \mathbb{R}^{r_q+s_q} \rightarrow \mathbb{R}$ , operating on a set of  $r_q$  past and present system inputs, and  $s_q$  past outputs. The LTIIP *observation model*,  $\mathbb{O}_{\theta_*, \varphi_*}$ , for  $t \in \mathbb{Z}$ , is given by

$$\begin{aligned} \mathbb{O}_{\theta_*, \varphi_*} : \quad y[t] &= \sum_{q=1}^Q \theta_{q*} \varphi_{q*}(x_{-\infty}^t, y_{-\infty}^{t-1}) + e_*[t] \\ &\stackrel{\text{def}}{=} \theta_*^T \varphi_* (x_{-\infty}^t, y_{-\infty}^{t-1}) + e_*[t], \end{aligned} \quad (3)$$

with  $\theta_* \in \mathfrak{P}$ , and  $e_* \in \mathbb{R}^Z$  an error sequence (characterized below) representing uncertainties in the model. The “\*” subscript indicates a “true,” but unknown, quantity associated with the observation model.<sup>3</sup> The arguments,  $x_{-\infty}^t$  and  $y_{-\infty}^{t-1}$ , of the regressor signals  $\varphi_q$  (or vector  $\varphi$ ) indicate

that a finite number of elements are selected from the subsequences  $\{\dots, x[t-1], x[t]\}$  and  $\{\dots, y[t-2], y[t-1]\}$  by each  $\varphi_q$  for processing at time  $t$ . For conservation of space, we define the vectors of signal samples,

$$\mathbf{u}_{q*}[t] \stackrel{\text{def}}{=} \begin{bmatrix} \text{vector of } r_q \text{ inputs from } x_{-\infty}^t, \\ \text{and } s_q \text{ outputs from } y_{-\infty}^{t-1}, \text{ used by} \\ \varphi_{q*} \text{ at time } t \end{bmatrix}, \quad (4)$$

and the matrix  $\mathbf{U}_*[t] = [\mathbf{u}_{1*}[t] \ \mathbf{u}_{2*}[t] \ \dots \ \mathbf{u}_{Q*}[t]]$ . Given observations of  $x$  and  $y$  sufficient to compute outputs on time interval  $t = 1, 2, \dots, \mathfrak{T}$ , we pose an *estimation model* as a function of the parameters and regressor signals,

$$\mathbb{M}_{\theta, \varphi} : y^p(t, \theta, \varphi) = \sum_{q=1}^Q \theta_q \varphi_q(\mathbf{u}_q[t]) \stackrel{\text{def}}{=} \theta^T \varphi(\mathbf{U}[t]), \quad (5)$$

in which each  $\varphi_q$  is drawn from the set  $\Xi_\varphi$  (see footnote 3),  $\theta \in \mathfrak{P}$ , and the  $\mathbf{u}_q[t]$  and  $\mathbf{U}[t]$  are defined similarly to (4). The superscript on  $y^p$  connotes “prediction,” as this estimation model corresponds to the classical prediction-error method of Ljung (1999). If the observation equation is modeled stochastically, and  $\mathcal{E}\{e_*[t]\} = 0$  for each  $t \in \mathbb{Z}$ , where  $\mathcal{E}\{\cdot\}$  denotes the expected value, then

$$\mathcal{E}\{y[t] \mid \theta_*, \varphi_*\} = \theta_*^T \varphi_*(\mathbf{u}_*[t]) = y^p(t \mid \theta_*, \varphi_*), \quad (6)$$

where the “conditioning” notation  $(\dots \mid \theta_*, \varphi_*)$  is used in a deterministic function merely to emphasize that the conditioning quantities are to be treated as fixed values. Thus, given “true” parameters and regressors, the prediction sequence will represent the minimum mean-square-error (MMSE) estimate of the observed sequence. That is,  $y^p(t, \theta, \varphi)$  is the MMSE predictor of the sequence produced by  $\mathbb{O}_{\theta_*, \varphi_*}$ . The prediction residual sequence associated with the general estimation model  $\mathbb{M}_{\theta, \varphi}$  is

$$\begin{aligned} \varepsilon(t, \theta, \varphi) &= y[t] - y^p(t, \theta, \varphi) \\ &= (\theta_* - \theta)^T \varphi_*(\mathbf{U}_*[t]) + e_*[t] \\ &\quad + \theta^T [\varphi_*(\mathbf{U}_*[t]) - \varphi(\mathbf{U}[t])], \end{aligned} \quad (7)$$

indicating error components due to the possible misadjustment in the parameter estimates as well as the possible improper selection of regressor functions. For “true” parameters and regressor signals, eq. (7) reduces to

$$\varepsilon(t \mid \theta = \theta_*, \varphi = \varphi_*) = e_*[t], \quad (8)$$

the asymptotic best-case residual.

To concomitantly find the model structure and parameter estimates, the regressor set is chosen according to an evolutionary view of the selection process. The parameters are identified using the set-theoretic approach which supports evolution by contributing model fitness measures.

## 3. SET-THEORETIC EVOLUTIONARY ALGORITHM

### 3.1 Evolutionary Model Selection

In the present formulation, a LTIIP model is an “*organism*” with a single “*chromosome*.” In contrast to the complex chemical structure of chromosomes in living cells, the LTIIP chromosome is a simple, finite, binary sequence in which each bit indicates the presence or absence of a particular “*gene*.” A binary “1” represents a region of “*coding DNA*” along the model chromosome, while a “0” represents “*non-coding DNA*.” Whereas the information in a biological gene is encoded in its sequence of nucleotides,

<sup>3</sup> The index  $q$  enumerates the elements of the set  $\Xi_\varphi$ , so the functions in (3) should be indexed as  $\varphi_{q_i*}, i = 1, \dots, Q$ , but we use the slightly abusive notation  $\varphi_{q*}$  for simplicity. It is to be understood that  $\varphi_{q*}$  is the  $q^{\text{th}}$  element selected from  $\Xi_\varphi$ , rather than element  $q$  of  $\Xi_\varphi$ .

Download English Version:

<https://daneshyari.com/en/article/711170>

Download Persian Version:

<https://daneshyari.com/article/711170>

[Daneshyari.com](https://daneshyari.com)